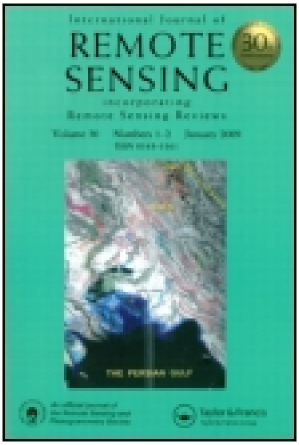

International Journal of Remote Sensing
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tres20

Monitoring Brazilian soybean production using NOAA/AVHRR based vegetation condition indices

W. T. Liu & F. Kogan

Available online: 25 Nov 2010

INT. J. REMOTE SENSING, 2002, VOL. 23, NO. 6, 1161–1179

Monitoring Brazilian soybean production using NOAA/AVHRR based vegetation condition indices

W. T. LIU

Research, Extension and Graduate study Center, Dom Bosco Catholic University, Campo Grande, 79117-900-MS, Brazil; e-mail: will@ucdb.br

and F. KOGAN

Climatic Research and Application Division, NESDIS/NOAA, NOAA Science Center 5200, Auth Rd, Camp Springs, MD 27026, USA;
e-mail: fkogan@orbit1s.nesdis.noaa.gov

(*Received 8 June 1999; in final form 9 August 2001*)

Abstract. This study explores the potential application of NOAA/AVHRR based satellite indices to estimate the soybean yield for Brazil. NOAA AVHRR GVI (Global Vegetation Index) weekly maximum composite NDVI (Normalized Difference Vegetation Index) data sets with a resolution of 16 km for the period of 1985 to 1998 (except 1995 due to missing data) provided by NOAA/NESDIS were used in this study. Nine soybean yield models, including eight principal production states and the country, were constructed using observed yield data and NDVI and/or TCI (Temperature Condition Index) data from 1985 to 1995. The data period of 1996 to 1998 was used to evaluate the model performance. The crop yield is generally affected by technological improvements through time and by the annual weather fluctuation. The contribution of technological improvements was approximated by trend term and weather-related fluctuations of yield around the trend were estimated through AVHRR-based indices. The results showed that four states had significant technological trend contribution with slopes ranging from 0.49 to 0.86 and R2 ranging from 0.22 to 0.62. In most of the models, yield variation around the trend was sensitive to TCI (Temperature Condition Index), during the period of the grain filling stage (end of January for northern states and mid February for the southern states). For most of the models, the determination coefficient (R^2) was higher than 0.6 and the root mean square error (RMSE) was lower than 10%. The results of model validation showed that the absolute errors were lower than 10% in 21 out of 27 cases tested. It is concluded that the satellite indices are useful for crop production monitoring. In the Brazilian soybean production region, where the summer crop season coincides with the rainy season, the temperature-based index is more informative about possible fluctuation of soybean yield and production in Brazil. It is suggested that a combination use of satellite and in situ data may likely improve the yield estimate.

1. Introduction

Brazil, with its 8.5 million square kilometres of territory, is the largest country of the South American continent, occupying nearly half of its land. Agriculture is

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online © 2002 Taylor & Francis Ltd
http://www.tandf.co.uk/journals
DOI: 10.1080/01431160110076126

the major branch of Brazilian economy and provides significant contribution to the gross national product. Agricultural exports of 16.4 billion dollars in 1997 account for about 30 per cent of the foreign exchange earned. The soybean export value of 2.4 billion dollars counts for 15% of the total agricultural export value in 1997. Brazil leads in South America and even the world in production of several crops, and many of them are vitally important for providing adequate amounts of food and feed. Among these crops, the soybean crop is a very important protein source for the human diet and for raising livestock.

Brazilian soybean production has risen from less than one million metric tons in 1968 (1% of world production) to 31.5 million metric tons in 1998 (20% of world production) according to FAO (1999). Although the Brazilian climate is generally favorable in the major areas of crop cultivation, variation in crop production from one year to another is rather high. For example, in 1990/91, when weather was unfavorable, total soybean production in Brazil dropped to less than 15 million metric tons, while in 1997/98, favorable weather stimulated good harvest resulted in 31.5 million metric tons of soybean. Therefore, crop monitoring and early assessment of final production in response to weather fluctuations are very important issues concerning government and traders in Brazil.

Weather data is a fairly good source of information used traditionally for monitoring crop growth and assessment of production. However, poor spatial distribution of weather stations often makes this task difficult to fulfill. Sometimes weather data is incomplete and/or not available early enough for timely assessments. In addition, weather observations are location-specific and do not adequately represent diversity of weather over the large areas where crops grow. Furthermore, assessment of final crop production requires estimation of cumulative environmental impact, which is only possible at the end of the harvest. Nevertheless, an earlier crop yield prediction with good precision before harvest is very important for optimizing its market value. Therefore, many studies are seeking different approaches for developing crop yield models (McQuigg 1975, Hodges *et al.* 1987).

Remote sensing data provides systematically high-quality spatial and temporal information about land surface features, including behaviour of agricultural crops and cumulative environmental impacts on crop growing conditions. It has been shown that AVHRR data obtained from NOAA polar orbiting satellite platforms can be used as a sole source of information, and can also be used complimentary to weather data, for the purposes of monitoring crop conditions and productivity on a large scale area (Tucker and Sellers 1986, Kogan 1990, 1995, 1997, Benedetti and Rossini 1993, Groten 1993, Quarmby *et al.* 1993, Hayas and Decker 1996, Rasmunsen 1997). Recently, the AVHRR data based three-channel numerical algorithm developed by Kogan (1997) was used to generate the satellite indices, including TCI and VCI (Vegetation Condition Index) for drought and crop production monitoring. This algorithm has been validated in Africa and showed promising results in early drought detection, vegetation growth condition monitoring, and crop yield assessment (Unganai and Kogan 1998). This research, in addition to focusing on a vitally important agricultural region, has the following objectives:

(i) to test the new AVHRR-based algorithm as an environmental indicator of vegetation condition in areas with sufficient moisture supply;
(ii) to investigate the nature of the relationship between satellite-derived and ground information based agricultural indicators; and

(iii) to simulate productivity of soybean in Brazil using new AVHRR-based indices as predictors.

2. Methods

2.1. *Areas of study*

The main soybean production regions are concentrated in the southern half of Brazil, including the southern states Rio Grande do Sul (RS), Paraná (PR) and Santa Catarina (SC); centre-western states Mato Grosso (MT), Mato Grosso do Sul (MS) and Goiás (GO); and south-eastern states Minas Gerais (MG) and São Paulo (SP). Table 1 and figure 1 present locations, areas where satellite data were collected and environmental details of these states. The eight states add up to 94% of the national total production in 1997. Administrative regions used in this study are large, cover different environmental and geographic zones and may not reflect homogeneous climatic conditions for a certain crop growing stage. Therefore, nine models were constructed in this study, including one model for each of the eight states counting for regional difference and one general country model (BR) taking the country as a whole.

2.2. *Environmental conditions*

In the soybean production regions the climate and vegetation type range from tropical wet–dry savanna to subtropical pine forest. The total annual rainfall ranges from 1200 mm to 1700 mm and annual mean temperature ranges from 18° C to 23° C.

The southern region of Brazil produces about 45% of the total national soybean crop. Its terrain ranges from vegetation-covered river valleys to rolling grasslands and forested highlands. The major market crops are rice, soybean, wheat, grapes, peach and apples. The climate in this region is classified as 'subtropical humid', with a total annual rainfall ranging from 1000 mm to 1650 mm, and with rainfall distribution varying from well distributed through out the year in the south, to gradually concentrated in the summer in the north. In general, the total rainfall during the summer crop season (October to March) is over 1000 mm, which indicates a moisture surplus when compared with the potential evapotranspiration for the same period (figure 1). The monthly mean temperature during the soybean crop season from

Table 1. The areas studied, with environmental information.

Region	Latitude (°S)	Longitude (°W)	Total annual rainfall (mm)	Mean annual temperature (°C)	Climate region	Vegetation type
BR	21.0° S–29.0° S	52.0° W–54.5° W	1200–1700	18–23	Tropical Subtropical	Savanna Pine Forest
MT	13.0° S–15.0° S	54.0° W–57.0° W	1200–1400	23	Tropical	Savanna
GO	16.5° S–18.0° S	48.5° W–52.0° W	1400–1600	23	Tropical	Savanna
MS	21.0° S–23.0° S	52.5° W–55.5° W	1400–1600	21	Tropical	Savanna
PR	23.0° S–25.0° S	51.0° W–54.0° W	1500–1700	19	Subtropical	Pine Forest
SC	26.5° S–27.0° S	52.5° W–54.0° W	1500–1700	18	Subtropical	Pine Forest
RS	27.0° S–29.0° S	53.0° W–55.6° W	1500–1700	19	Subtropical	Grass land
SP	21.0° S–23.0° S	50.0° W–52.5° W	1300–1500	20	Tropical	Savanna
MG	18.5° S–20.0° S	47.0° W–51.0° W	1200–1400	21	Tropical	Savanna

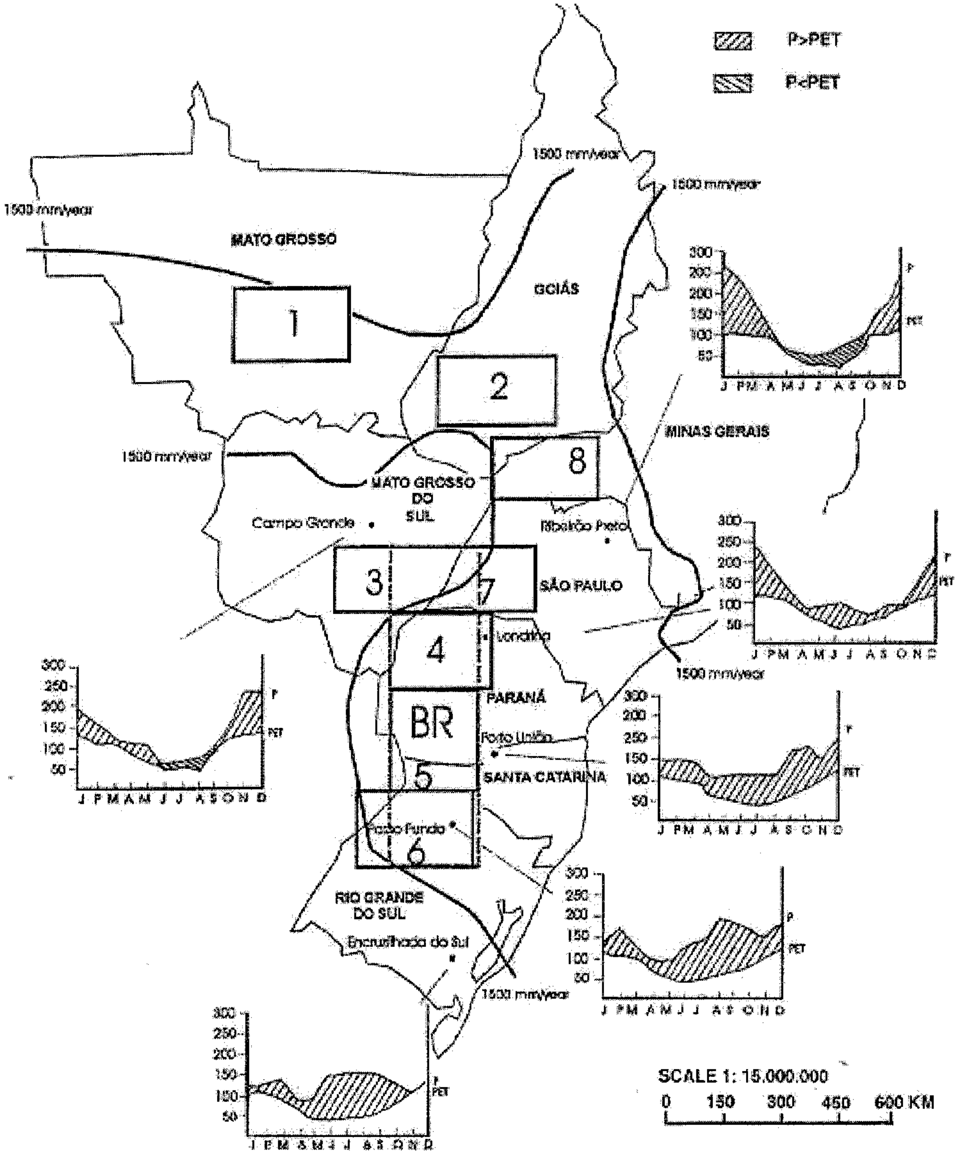

Figure 1. Administrative regions, annual mean precipitation isoline (PCP, in mm), monthly averaged precipitation (mm) and monthly mean potential evapotranspiration (PET, in mm) for selected stations (Source: USDA, 1994).

October to March is mostly over 20°C and the maximum temperature in this period is mostly over 26° C.

The centre west region of Brazil is a rolling Savanna called the 'Cerrados'. Due to government incentives, farmers migrated from the south in the early 1980s, introducing updated agriculture management technologies. In the past 12 years, this region has seen a rapid expansion of soybean cultivation acreage (table 2). The region has distinct wet–dry seasons, with monthly rainfall ranging from near 10 mm in the winter months (June to August) to over 200 mm in the summer cropping

Table 2. Areas where soybean is harvested, with production evolution between 1986 and 1997.

	Area (Mha)				Production (Mt)			
Region	1986	% of Brazil	1997	% of Brazil	1986	% of Brazil	1997	% of Brazil
BR	9.18	100	11.50	100	13.33	100	26.43	100
MT	0.91	10	2.19	19	1.92	14	6.06	23
GO	0.62	7	1.02	9	0.92	7	2.45	9
MS	1.20	13	0.88	8	1.96	15	2.18	8
PR	1.70	18	2.54	22	2.60	20	6.57	25
SC	0.38	4	0.23	2	0.50	4	0.54	2
RS	3.20	35	2.89	25	3.27	24	4.66	18
SP	0.48	5	0.57	5	0.92	7	1.41	5
MG	0.44	5	0.50	4	0.80	6	1.10	4
Total	8.93	97	10.80	94	12.89	97	24.95	94

season (December to February). The monthly mean temperature remains close to 25° C during summer cropping season. Figure 1 shows the monthly mean rainfall (P) and monthly potential evapotranspiration (PET) of three stations, Cuiabá, Gioánia and Campo Grande, located in this region. The maximum temperature in the cropping season is mostly higher than 28° C.

The south-eastern region of Brazil is the most heavily industrialized area of the country and contains Brazil's three largest cities: São Paulo, Rio de Janeiro and Belo Horizonte. Characterized by rolling, hilly uplands, mainly the Cerrados with deep red soils, the region provides favourable conditions for agriculture development. The region is dominated by two states: São Paulo and Minas Gerais. São Paulo has the most diverse agriculture, including sugar cane, orange, soybean, cotton, corn, coffee and livestock. Most of this produce is a significant component of the total national production. The soybean production in this region counts for 7% of the national total. The cultivation areas are concentrated in the north of São Paulo and in the north-west of Minas Gerais. The region also has a distinct wet–dry season with monthly rainfall ranging from close to 20 mm in the winter months to over 200 mm in summer peak rainy season. The monthly mean temperature is around 23° C and the maximum temperature is around 26° C in the summer cropping season.

In all regions, monthly mean rainfall exceeds monthly mean potential evapotranspiration during the summer cropping season (figure 1). This indicates that in general there should be water surplus for annual crops. Nevertheless, there are frequent occurrences of dry spells during the peak growing stage in January and February, which often affect the annual crop yield.

2.3. *Soybean data*

Soybean yield data and NOAA/AVHRR data for the period of July 1985 to June 1998 were used in this study. Eight states (large administrative regions) where soybean is an important crop in the regional economy were selected for analysis; and state-average soybean yield data provided by the Brazilian Geography and Statistics Institute (IBGE, 'Instituto Brasileiro de Geografia e Estat´sticos') were used in this study. Country soybean yield data were collected to study the potential application of satellite data for extremely large area estimation.

Area sampling methods and aggregation of crop information from county to state and country level are routinely used by IBGE for reporting crop production in Brazil. This government agency, responsible for national crop data collection, routinely gathers crop data, including planted and harvested areas, crop growing conditions, yield and production, through local IBGE agencies. These official yield statistics reported by IBGE are considered as a standard source of information by the Brazilian government, and hence are used in this study as the observed crop data (IBGE 1999). The soybean yield data for the period of 1986 to 1995 were used for model construction while those during 1996–1998 were used for model validation.

2.4. *Satellite data*

A NOAA AVHRR GVI (Global Vegetation Index) data set provided by NOAA NESDIS (National Environmental Satellite Data and Information Service), was used in this study. The spatial resolution of NDVI data is 16 km by 16 km at the equator and gradually degrades to 20 km by 20 km at both poles. Radiance values of three AVHRR channels, including visible (0.58–0.68 μm, Ch1), near infrared (0.72–1.1 μm, Ch2), and thermal (10.3–11.3 μm, Ch4) were used. The Ch1 and Ch2 radiances, were calibrated following Rao and Chen (1995, 1996) and the Normalized Difference Vegetation Index ($NDVI=(Ch2-Ch1)/(Ch2+Ch1)$) was calculated. The Ch4 radiance was converted to brightness temperature (BT) using look-up table and the values obtained were corrected in order to remove errors related to non-linear behavior of the sensor (Weinreb *et al.* 1990, Kidwell 1996).

AVHRR data were extracted for a South America continent window from 13° N to 30° S. This sample window is large and covers different environmental zones, which have uneven distribution of soybean plantations throughout. Therefore, to aggregate regional average values of satellite indices, data were collected only from areas of predominant soybean concentration (IBGE 1999). Coordinates of the data collection areas are shown in table 1. For the satellite data to match the southern hemisphere meteorological and crop growing cycle, the soybean crop calendar year was reconstructed to stretch from July of the current year through to June of the following calendar year. Week 1, therefore, refers to the first week of July, week 26 to the last week of December and week 52 to the last week of June in the following calendar year.

2.5. *Vegetation condition indices*

The three-channel algorithm comprehensively processes of NDVI and BT, including complete removal of temporal high frequency noise, stratification of world ecosystems, and detection of low frequency fluctuations associated with weather-induced changes in vegetation condition (Kogan 1995, 1997). These steps are crucial in order to obtain index data suitable for analysis and interpretation of vegetation growth conditions and regional environmental impacts.

The currently available physical and logical algorithms are unable to remove high frequency noise. Therefore, an alternative statistical technique, '4253H-twice', proposed by Van Dijk *et al.* (1987) was used to remove high frequency signals caused by hot spots and low frequency signals caused by cloud contamination in NDVI and BT time series. In order to preserve the identity of local conditions each pixel was processed separately. After smoothing, weather-dependent differences in NDVI and BT time series become more apparent. However, further enhancement is required because the component of variation induced by the weather is considerably smaller in magnitude than that produced by the ecosystem. For example, in highly productive grassland in the USA (Illinois) the range of NDVI variation between years with extreme weather conditions is less than one half of the typical value for that ecosystem (Kogan 1995). Conversely, the NDVI in north-eastern Brazil, at a specific pixel during the peak rainy season, varies 3–4-fold between a year with severe drought and a year with good rainfall (Liu and Kogan 1996). Therefore, an important step in algorithm development is to separate weather and ecosystem components (Kogan 1995).

This separation involves stratification of the ecosystems by calculating historical minimum and maximum of NDVI and BT at a certain pixel. The multi-year maximum NDVI characterizes the highest greenness value in that pixel. Most years, the planting date of soybean crop in Brazil starts at the beginning of the rainfall season, which is from the end of October to the beginning of November. By considering that soybean crop growing stages do not vary greatly each year, this maximum NDVI value is then used to infer the highest crop growth vigour of a certain growth stage at a certain pixel. It is then assumed that the historical maximum crop yield is highly correlated with the sum of all the maximum NDVI values through out the crop growth cycle. In contrast, the minimum NDVI characterizes the lowest crop growth vigour of a certain growth stage at a certain pixel; and the historical lowest crop yield is assumed to be highly correlated with the sum of all the minimum NDVI throughout the crop growth cycle. As the air temperature in a higher greenness crop canopy is known to be generally lower than the air temperature immediately above it, the maximum BT is defined conversely to the maximum NDVI. This means that a higher BT value indicates unfavourable crop growth conditions, whereas a higher NDVI indicates better growth conditions. Therefore, the absolute maximum and minimum of NDVI and BT calculated from ten years of historical satellite data were used as criterion to calculate crop growth condition at each week through out the years studied (Kogan 1995).

Given these considerations, the largest and the smallest NDVI and BT values of each pixel during 1985–1994 were calculated for each of the 52 weeks throughout a year. They were used as the criteria for determining the upper and the lower limits of the potential crop yield. These limits characterize the yield capacity (the range of NDVI and BT fluctuation) in response to year-to-year weather variability in each pixel. Furthermore, satellite indices, by taking weekly relative values of NDVI and BT with respect to their maximum and minimum values, were calculated for each pixel and for each year. Equations 1 and 2 describe the calculation of VCI and TCI indices. Since NDVI and BT interpret opposite extreme weather events, a lower BT was considered to indicate better crop growth conditions. Therefore, a higher TCI value indicates a lower BT. Equation 2 was modified to reflect this response of the crop to low BT. Hence, a higher TCI value indicates favourable crop growth conditions, which agrees with a higher greenness inferred by a higher NDVI value.

$$\mathrm{VCI}=\frac{(\mathrm{NDVI}-\mathrm{NDVI}_{\mathrm{min}})}{(\mathrm{NDVI}_{\mathrm{max}}-\mathrm{NDVI}_{\mathrm{min}})}\times 100\% \tag{1}$$

$$\mathrm{TCI}=\frac{(\mathrm{BT}_{\mathrm{max}}-\mathrm{BT})}{(\mathrm{BT}_{\mathrm{max}}-\mathrm{BT}_{\mathrm{min}})}\times 100\% \tag{2}$$

where NDVI, $\mathrm{NDVI}_{\mathrm{max}}$ and $\mathrm{NDVI}_{\mathrm{min}}$ are smoothed weekly NDVI, multi-year absolute maximum, and multi-year absolute minimum, respectively; and BT, BT_{max}, and BT_{min} are similar values for brightness temperature.

2.6. *Trend yield*

Crop yield is generally affected by technological improvements through time and by the annual weather fluctuation. In general, the productivity of a certain crop in a certain region increases with cultivation history. This increase is a sum of human efforts to improve productivity by selecting for varieties with higher productivity and higher resistance to environmental stresses; and also improved crop management technologies. Crop management technologies include mechanization, irrigation, fertilization, soil and water conservation, cropping system and controls of weeds, insects and diseases. The combination of these factors for crop improvement is known as the historical technology trend term (McQuigg 1975). In this study, the contribution of technological improvements to crop yield was approximated by a trend term expressing the historical soybean yield as a function of time. The first step of constructing a crop yield model is to find out the trend yield and separate its influence from the influence of the weather. In this study, the trend yield is defined as the annual yield increment as a function of year increment using the data period of 1986 to 1995. In order to illustrate the results with figures in the same magnitude in this study, the crop yield unit in kg/ha was converted to 100 kg/ha, and the calendar year unit was converted to 100 years. Therefore, the trend yield is a linear function obtained by correlating the observed yield to its calendar year, as follows:

$$Yt = a - by \tag{3}$$

where Yt = trend yield (100 kg/ha), a = interception coefficient, b = slope of observed yield to year, and y = crop calendar year (100 years).

2.7. *Yield departure model*

After the annual yield was separated from its trend yield, the model was constructed using the percentage of yield departure (dY) from its trend yield as a dependent variable and TCI and VCI as independent variables. The dY is defined as a percentage as follows:

$$dY = \frac{(\text{Observed yield})}{\text{Trend yield}} \times 100\% \tag{4}$$

Therefore, a multiple linear regression model, using dY as dependent variable and AVHRR-based indices TCI and VCI as independent variables, was constructed to estimate the contribution of weather-related fluctuations expressed by dY for each studied region. Model performance was evaluated, first by calculating the dY predicted by the model, and then applying the equations 3 and 4 to obtain the estimated yield. The yield unit was then corrected to its original unit of kg ha^{-1} by multiplying by 100.

3. Results and discussion

3.1. *Trend yield*

Brazil's relatively short history of soybean cultivation began when Taiwanese immigrants introduced it in the early 1960s. The total planting area was about one million hectares at the end of 1960, and was concentrated mostly in the Rio Grande do Sul and Paraná States. Contribution of the other states was significant only after the government's agricultural expansion plan, which was launched in the early 1970s. Nationwide, the planting area has increased approximately 20% in the past 12 years from 9.18 million hectares in 1986 to 11.5 million hectares in 1997, while the production has doubled from 13.33 to 26.43 million metric tons within the same period. In the Mato Grosso State, in particular, the planting area has expanded from a half million hectares in 1984 to around 2.65 million hectares in 1998. This state has the most significant soybean expansion in Brazil during the past fifteen years, increasing production from one million tons in 1984 to around 7 million tons in 1998. Conversely, in the southern states, including Paraná, Santa Catarina and Rio Grande do Sul, with their longer history of soybean cultivation, planting area and production have stabilized and even showed signs of decreasing. Table 3 presents technology-based yield evolution during the period of 1986 to 1995. In general, during this period, most of the soybean production regions showed a considerable yield increase: 52% average increase across the country, ranging from a 4% increase in the Goiás state to as high as 93% in Rio Grande do Sul.

The annual increase rates of yield due to technological improvements in each region were obtained by applying the linear regression technique, using historical yield data against year. Table 4 presents the estimated technological trend term from 1986 to 1995. Within the nine regions shown, MT, PR, SC, RS and BR had a significant trend term contribution, with a regression slope of 0.49 to 0.86. A high slope of 0.49 ($R^2 = 0.62$) for MT indicated that, in addition to a rapid area expansion, there was considerable yield growth due to technological improvements (table 2). Although a high slope of 0.86 was also obtained for the SC, the state decreased its planting area from 0.437 million hectares in 1984 to 0.28 million hectares in 1993. Because the main soybean plantations are concentrated in the western highland subtropical pine forest, where in recent years other economically valuable crops were gradually replacing the soybean crop. The decline of planting area and the positive trend may indicate that farmers have to adapt new crop management technologies

Table 3. Variations of soybean yield due to technological improvements and weather impact, 1986–1995.

	Technology			Weather	
	Trend yield ($100\,kg\,ha^{-1}$)			Yield departure from trend** (%)	
Region	1986	1995	Yield change* (%)	AMD***	Variation range
BR	14.50	21.99	52	6.70	−19–13
MT	20.82	23.64	14	4.04	−12–6
GO	18.47	19.14	4	7.37	−33–9
MS	19.57	21.88	12	7.36	−18–12
PR	14.90	25.57	72	9.96	−21–17
SC	13.02	21.77	40	9.16	−42–10
RS	10.08	19.45	93	19.26	−23–17
SP	19.29	22.37	16	6.54	−16–15
MG	18.15	19.97	10	6.22	−31–7

*Yield change = 100(yield of 95 − yield of 86)/yield of 86.
**Yield departure from trend = 100(observed yield)/trend yield.
***AMD = absolute mean yield departure from trend.

Table 4. Estimate of the soybean yield trend during 1986 to 1995 (yield data source: IBGE).

Region	Regression intercept	Coefficient slope	R^2 (%)	RMSE (%)
MT	−21.43	0.4901	0.62	6
GO	1.60	0.1965	0.08	18
MS	−6.86	0.2973	0.23	11
PR	−27.25	0.5364	0.34	15
SC	−61.89	0.8608	0.57	26
RS	−38.87	0.6038	0.22	48
SP	2.52	0.1942	0.12	10
MG	2.81	0.1837	0.08	16
BR	−40.86	0.6593	0.58	10

in order to increase soybean productivity and make it more competitive. The remaining states had R^2 values lower than 0.22 and slopes lower than 0.2, which implied that technology contributed little to increased soybean yield in 1986–1995.

3.2. *Correlation of soybean yield to TCI and VCI*

From table 3, it can be seen that a high variation in absolute mean yield departure from trend (AMD) ranged from 4.04% in MT to 19.26% in RS. Absolute departures ranged from −42% in SC to +17% in PR. These observations explained that, besides the contribution of technological trend to yield increase, the annual weather variability significantly affects soybean yield in Brazil.

The annual crops in Brazil, such as soybean, rice, corn and cotton are mostly planted at the beginning of summer rainy season (October to November) and harvested at the end of rainy season (March to April). The major variety of soybean planted in Brazil has a phenological cycle of 125 days. The annual total rainfall in soybean regions ranges from 1200 mm to 1500 mm. However, during January and February, dry spells, so-called 'veranico', are observed. The period when soybean is highly sensitive to weather is from November to March. This period was investigated here for model development.

It is reasonable to assume that the farmers start their seeding once the soil has favourable moisture conditions. There should be no crop loss at the early growing stage. Therefore this early growing period, which is approximately from week 16 to week 20, corresponding to the third week of October to the first week of November, was not used for model construction. In this study, the correlation between dY and weekly TCI, and between dY and weekly VCI for the soybean growth period of week 21–40, corresponding to the third week of November to the second week of April, for each region, are shown in figure 2. For the regions MT, GO, MS, SP, MG and BR, the highest correlation of dY against TCI occurred between week 30 (the end of December) and week 33 (the third week of January). This period generally occurs during the flowering and grain filling phenological stages, which are more sensitive to climatic fluctuation. The highest correlation between dY and TCI occurred in week 28 for PR and in week 35 for RS. For SC, the correlations were all over 0.7 after week 31. In general, there is a clearly identified critical period during which dY is sensitive to TCI variation. Most of the correlation coefficients were high, ranging from 0.48 to 0.78. Figure 3 shows the dynamics of dY and TCI for the week with highest correlation. In most cases, TCI agreed well with dY. In

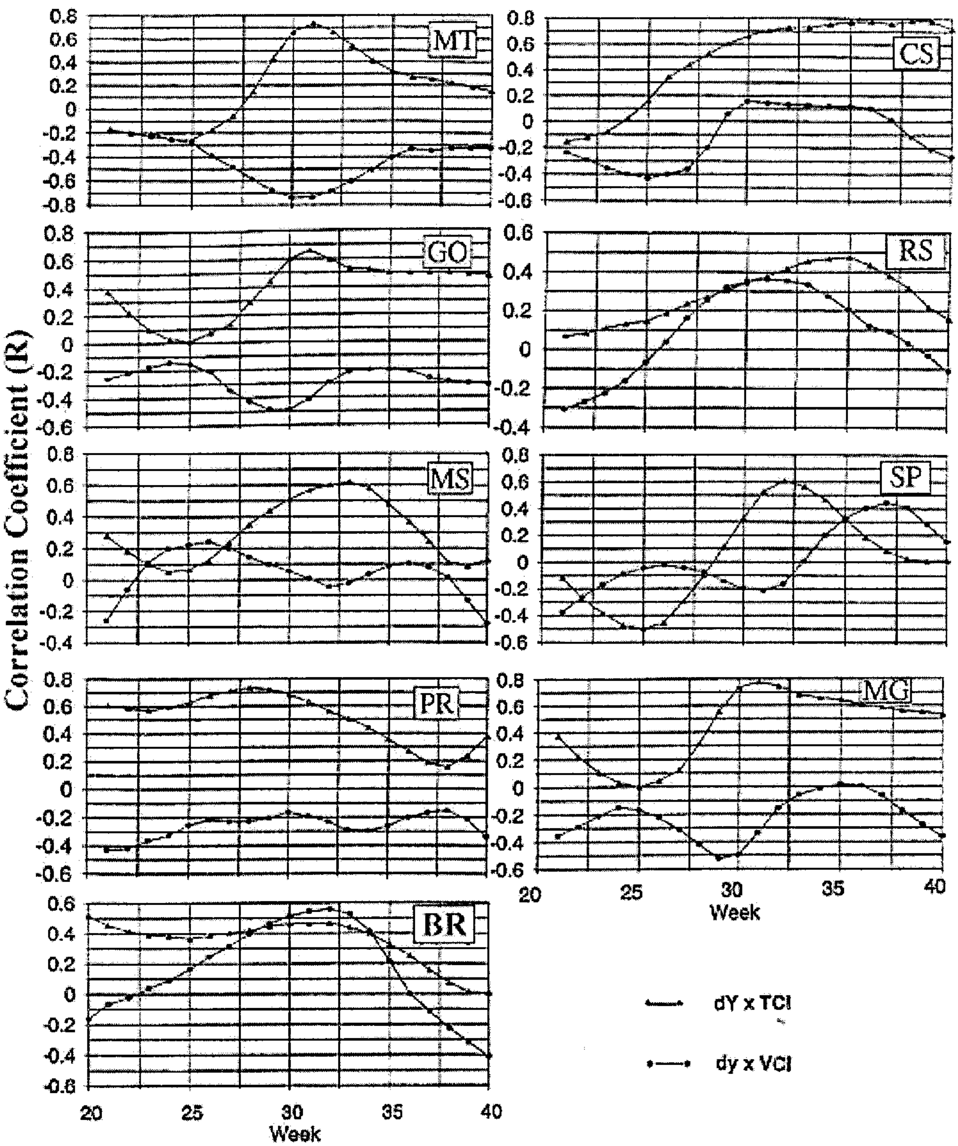

Figure 2. Time series plots of correlation coefficient (R) between yield departure from trend (dY) and TCI and between dY and VCI during the soybean crop growth period (week 20 to week 40 corresponding mid November to mid April).

1990 GO and MG had the highest yield departure from the trend yield line (−33% and −30%, respectively) and lowest TCI (11% and 4%, respectively). In 1991, the southern states, including PR, SC and RS, showed considerable yield departure from the trend yield line, which also coincided with the lowest TCI values. The spatial difference in TCI values between 1990 and 1991 for week 31 (29 January–4 February) is seen in figure 4(*a*). When comparing colour variation between years, it can be seen that red dominated south-eastern Brazil (south of 21° S) in 1990 and 1991, indicating thermal vegetation stress, while in 1997, blue dominated the same area, indicating favourable conditions. Regarding two drought years in southern Brazil, vegetation

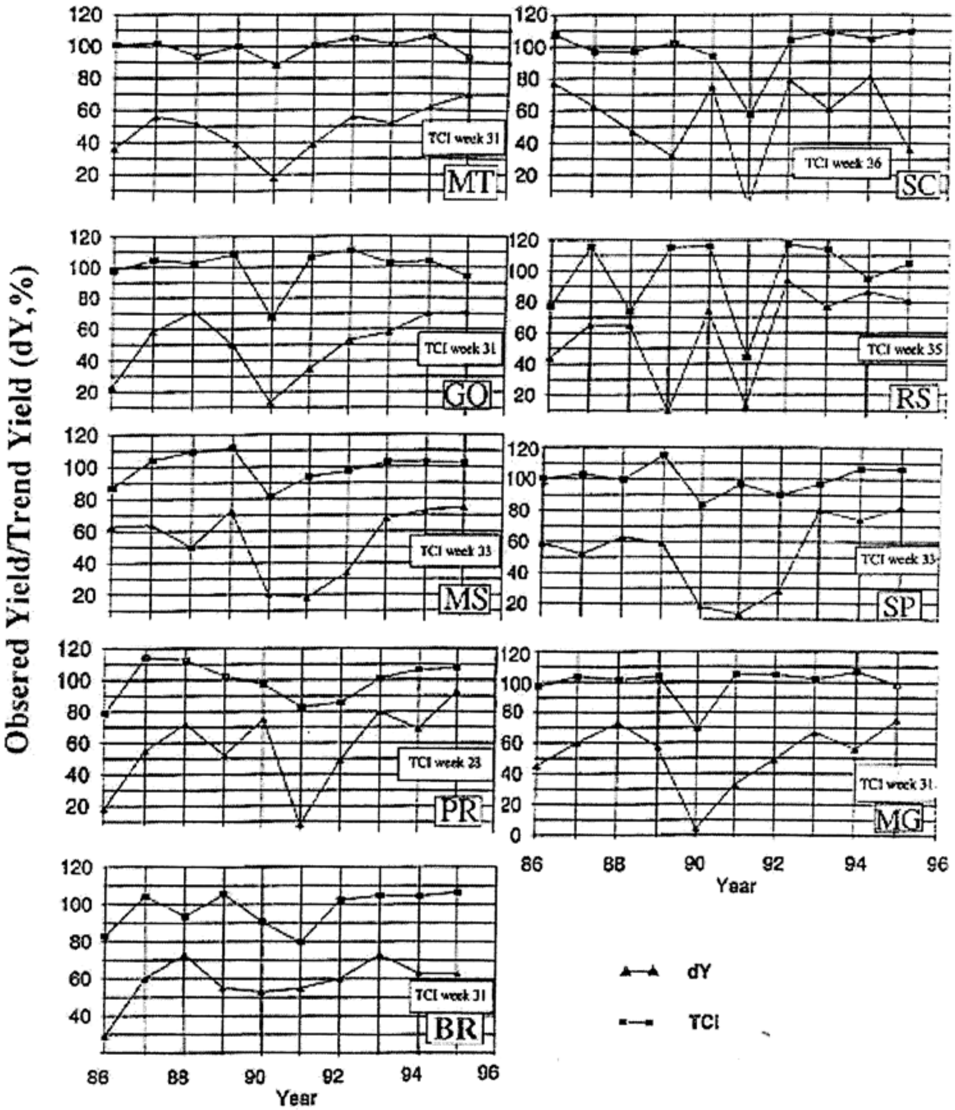

Figure 3. Comparison of soybean yield departure from trend (%) with the highest correlation week of TCI during 1986–1995.

stress was more severe in 1991 than in 1990 (regions PR, SS and RS). Conversely, east of 51° W, conditions in 1990 were much worse (with an intensive red colour) than in 1991 (with a mix of red, green and blue). This resulted in low 1990 soybean yields in MT, GO, MS, SP and MG. Nevertheless, the correlation of dY against VCI were general low. This indicates that the soybean yield is much less sensitive to VCI fluctuations except for in RS, SP and BR, where R values were slightly over 0.4 (figure 3). As seen in figure 4(*b*), there are only slight differences in VCI values between two drought years (1990 and 1991) and one favourable year (1997). In order to use VCI for model construction, a modified stress based on VCI, which counts a non-linear effect on yield loss, may improve the model performance.

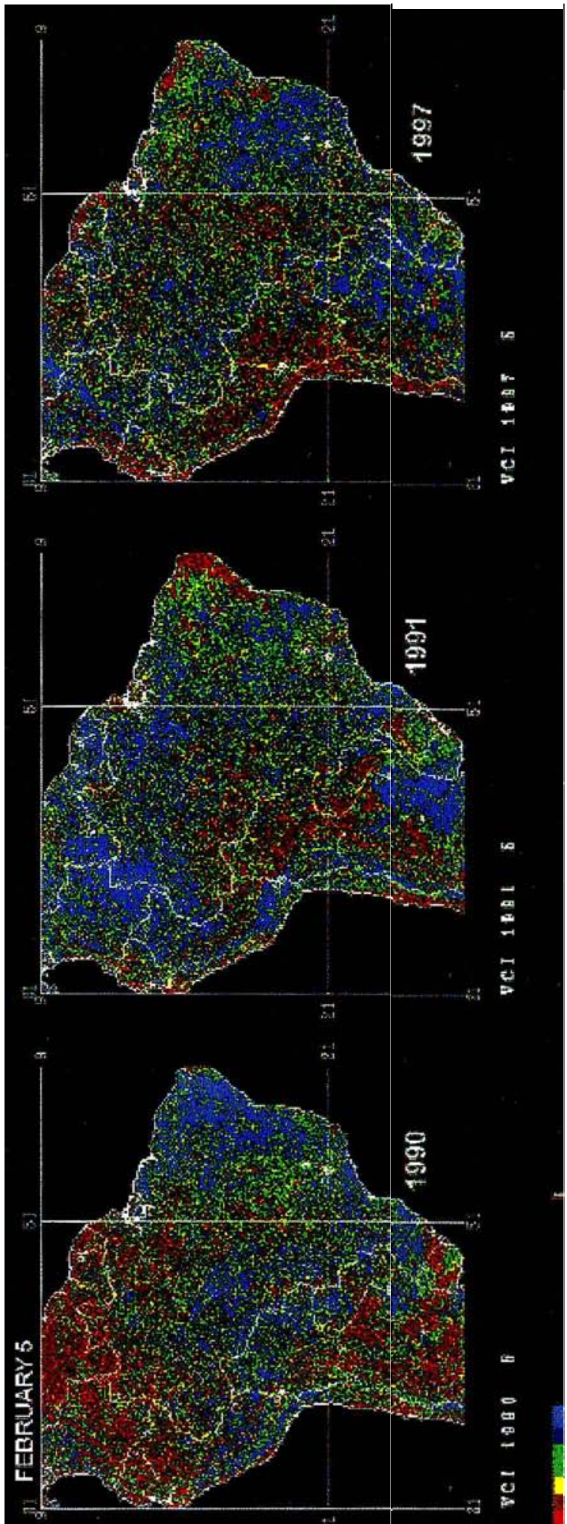

Figure 4. (*a*) Images of TCI week 28 (5–11 February) of 1990, 1991 and 1997, Brazil.

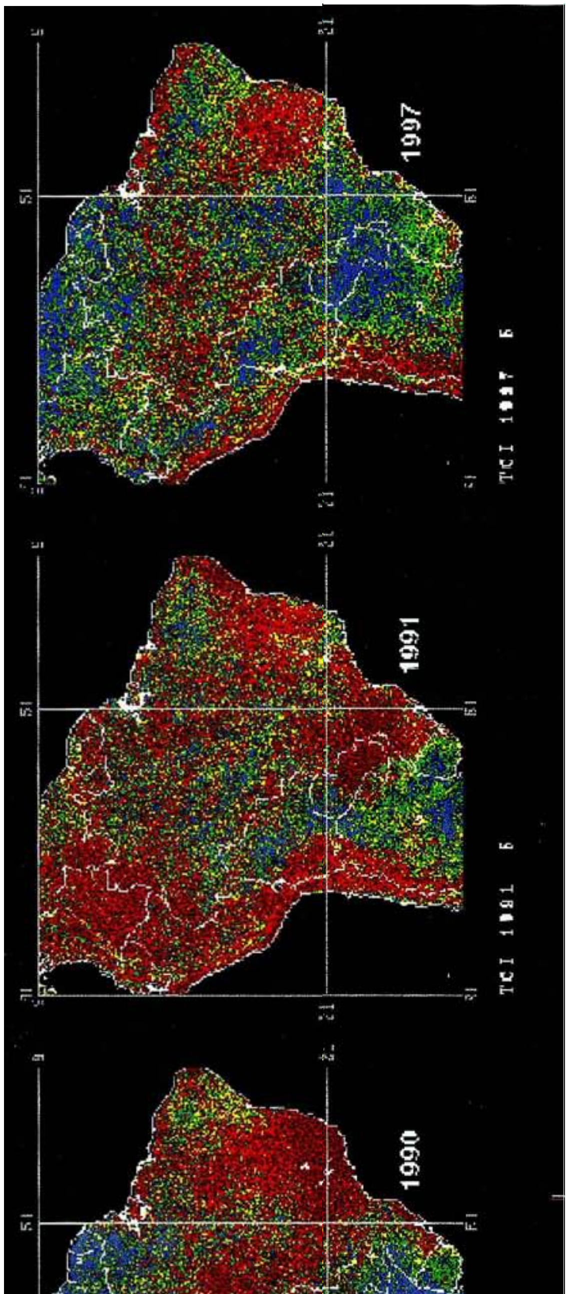

Figure 4. (*b*) Images of VCI week 28 (5–11 February) of 1990, 1991 and 1997, Brazil.

3.3. *Yield models*

The model is constructed in two steps: firstly modelling the trend yield, and then modelling dY. The trend yield model is linear, and includes an interception constant and a regression coefficient (slope). Table 4 in the trend yield section presents the trend yield models for the nine regions studied. For the dY models, dY data were used as dependent variables and the mean values of TCI and/or VCI during 1–4-week periods were used as independent variables. In order to select better candidate variables, correlation coefficients lower than 0.3 were excluded. One to two candidate models were then selected through the 'STEPWISE' regression process available with SAS software ('Statistical Analysis System'). Each model contains only two variables using the data period from 1986 to 1995. Data from 1996 to 1998 were used for model validation.

Table 5 summarizes the dY models. Most of the models had high R^2 values ranging from 0.54 to 0.92, except RS with 0.29 and BR with 0.35. The RMSE were less than 10% except for RS with 25%. In most models, TCI (VCI in some models) from January (week 30) to mid February (week 35) were good predictors of dY during soybean flowing to grain filling stages. Although the correlation coefficient of dY to VCI was rather low in most cases, there were four models, including MS, RS, SP and BR, where both TCI and VCI were used as independent variables.

3.4. *Model validation*

The final yield estimate was obtained by applying the TCI and/or VCI data required by each model described in table 5 to calculate dY and the yield estimate,

Table 5. Models and statistics for yield departure from trend (dY) as a function of TCI and/or VCI.

Region	State/Country	Model*	R^2	RMSE (%)
1	Mato Grosso	$82.484 + 0.51068T5 - 0.0026235(T6)^2$	0.71	4
2	Goiás	$43.437 + 2.61686T_{31} - 0.02568(T_{31})^2$	0.86	6
		$43.092 + 2.6005T3 - 0.02527(T3)^2$	0.81	6
3	Mato Grosso do Sul	$69.0754 + 0.3442T4 + 0.23195V4$	0.54	8
4	Paraná	$69.1393 + 0.9628T4 - 0.00656(T4)^2$	0.59	9
5	Santa Catarina	$56.5566 + 1.3685T3 - 0.00954(T3)^2$	0.88	6
6	Rio Grande do Sul	$59.88 + 0.3539(V5)^2 + 0.00387(T7)^2$	0.29	25
7	São Paulo	$71.4397 + 0.3572T6 + 0.0034(V7)^2$	0.86	4
8	Minas Gerais	$66.0302 + 1.405T_{32} - 0.01266(T_{32})^2$	0.92	4
		$65.3574 + 1.4499T2 - 0.01323(T2)^2$	0.92	4
9	Brazil	$76.1875 + 0.284V3 + 0.00096(T3)^2$	0.35	9

*Models: T represents TCI; V represents VCI; independent variables are shown below:

- For region 1: $T5 = (T_{30} + T_{31} + T_{32})/3$; $T6 = (T_{33} + T_{34} + T_{35})/3$.
- For region 2: $T3 = (T_{30} + T_{31} + T_{32})/3$.
- For region 3: $T4 = (T_{31} + T_{32} + T_{33} + T_{34})/4$; $V4 = (V_{24} + V_{25} + V_{26} + V_{27})/4$.
- For region 4: $T4 = (T_{26} + T_{27} + T_{28} + T_{29})/4$.
- For region 5: $T3 = (T_{35} + T_{36} + T_{37})/3$.
- For region 6: $T7 = (T_{36} + T_{37} + T_{38})/3$; $V5 = (V_{30} + V_{31} + V_{32})/3$.
- For region 7: $T6 = (T_{33} + T_{34} + T_{35})/3$; $V7 = (V_{36} + V_{37} + V_{38})/3$.
- For region 8: $T2 = (T_{31} + T_{32})/2$.
- For region 9: $T3 = (T_{31} + T_{32} + T_{33})/3$; $V3 = (V_{31} + V_{32} + V_{33})/3$.

Where the subscript number indicates the week number; the normal number indicates the period summed through the number of weeks, which varies from region to region.

by multiplying the trend yield value with dY using the trend yield model in table 4. The final yield in kg/ha was then obtained by multiplying by 100. Figure 5 shows simulated and observed yield for the period of 1985 to 1998 (except for 1995).

Within the period used for model construction (1986 to 1994), the simulated yield for MT, GO, MS, SP, and MG corresponded quite well with low yield in 1990, while the models for PR, SC and RS corresponded quite well with the low yield in 1991. Exceptions were estimates for MS and RS. These models had lower R^2 values (MS with 0.54 and RS with 0.29), compared with the others. The national model also had a low R^2 value (0.35), which resulted in larger errors in 1987, 1989 and 1991. Most of the models had RMSE values of lower than 10%. The climatic pattern of the soybean production region in RS indicates excessive wetness all year around. This excessive wetness occurring during maturity and harvest stages may contribute to a higher estimation error for RS, since the model did not include TVI/VCI in these two stages. It is suggested that several complimentary step-by-step models, which combine the weather forecasting through the whole crop growing stages, might improve the final yield estimate.

The models were then evaluated independently by comparing simulated and observed yields during 1996–1998. These yield data were not included in the model construction. Table 6 and figure 5 show the results of the comparison. Among nine

Table 6. Observed and simulated soybean yields (independent test).

State/Brazil	Year	Yield (kg/ha) Observed	Yield (kg/ha) Estimated	Error (%)
MT	1996	2550	2484	−3
	1997	2766	2852	3
	1998	2702	2618	−3
GO	1996	2270	1899	−16
	1997	2412	2240	−7
	1998	2467	2224	−10
MS	1996	2408	2456	2
	1997	2467	2496	1
	1998	2091	2484	18
PR	1996	2677	2443	−9
	1997	2580	2583	0
	1998	2558	2630	3
SC	1996	2438	2185	−10
	1997	2324	2266	−2
	1998	2355	2266	−4
RS	1996	1769	1907	8
	1997	1613	1870	16
	1998	2088	2494	19
SP	1996	2144	2320	8
	1997	2450	2342	−4
	1998	1948	2438	25
MG	1996	1953	2065	6
	1997	2201	2033	−8
	1998	2261	2188	−3
BR	1996	2195	2284	4
	1997	2298	2209	−4
	1998	2360	2490	6

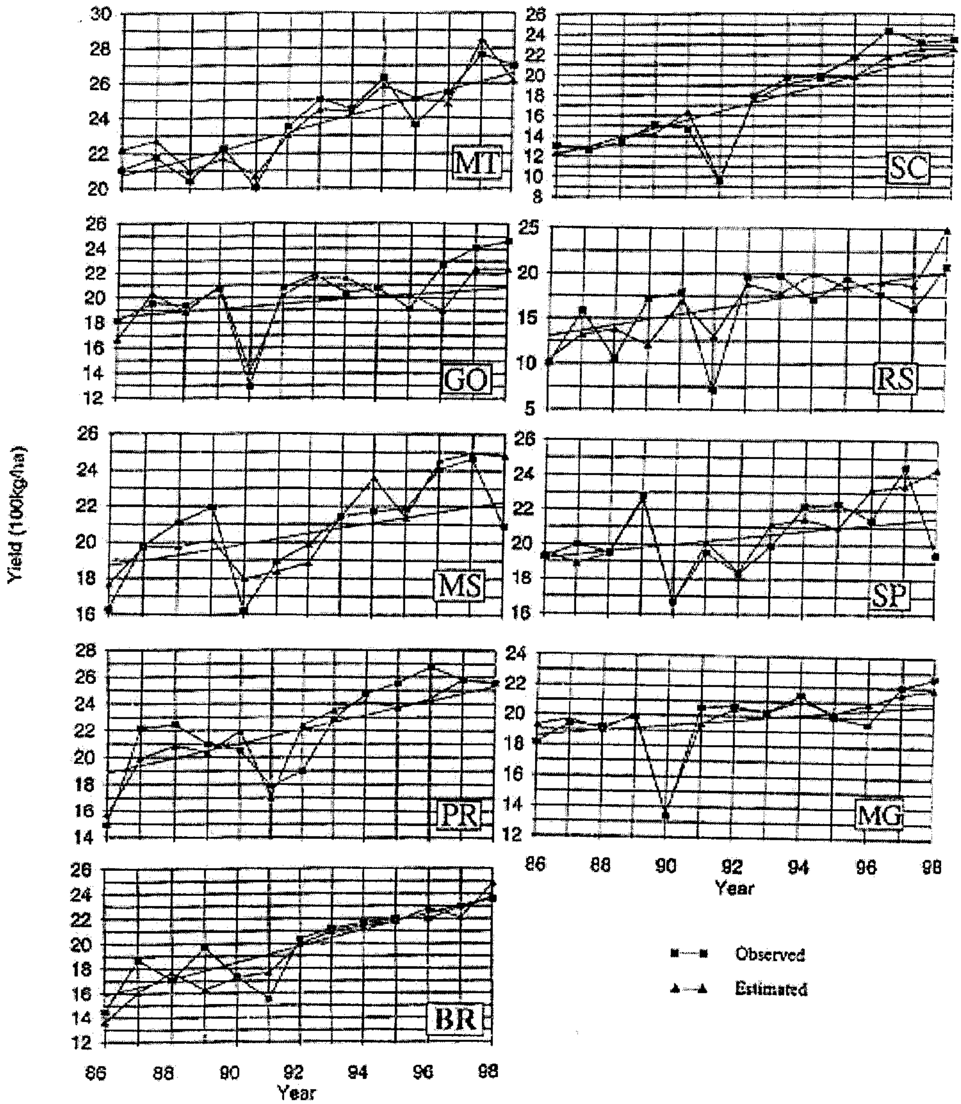

Figure 5. Comparison of the observed technological trend and estimated soybean yields during the period of 1986 to 1998 (excepting 1995).

tested models, five of them had an absolute error of 10% or less (regions MT, PR, MG, SC and BR). The SC model had an error of 10% but all underestimating. For the remaining models, in some years yield was overestimated by over 10%. The results of both dependent and independent model evaluations showed that the higher error obtained for RS was expected, since the RS model had low R^2. Other possible causes of error may be due to the climatic impact during the mature stage of soybean cultivation, a period which was not included in the model construction.

4. Conclusion

It is concluded that the AVHRR-based indices explored in this research are useful for crop production monitoring. In the major areas of Brazilian soybean cultivation, soybean yield prediction models were based on either TCI/VCI or TCI alone, which demonstrates the possibility of using TCI as well as VCI for developing soybean crop yield prediction models. In some regions, a combination of satellite and *in situ* data may be likely to improve yield estimate. It is suggested that using a combination of satellite and *in situ* data may improve the yield estimate. A future study will include the application of the same techniques with the incorporation of *in situ* data, to construct yield forecasting models of other important grain crops in Brazil such as corn, wheat, cotton and rice.

Acknowledgments

The authors wish to thank Dr D. Tarpley and Dr Jerry Sullivan for fruitful discussions and suggestions. This research work is the result of the research project financed by the FAPESP, São Paulo State, Brazil, under contact: Proc. N°. 1997/11568-3.

References

Benedetti, R., and Rossini, P., 1993, On the use of NDVI profiles as a tool for agricultural statistics: The case study of wheat yield estimate and forecast in Emilia Romagna. *Remote Sensing of Environment*, **45**, 311–326.

FAO, 1999, FAOSTAT, Agriculture Data, *Rome, Italy*. http://apps.fao.org/cgi-bin/nph-db.pl.

Groten, S. M. E., 1993, NDVI—crop monitoring and early yield assessment of Burkina Faso. *International Journal of Remote Sensing*, **14**, 1495–1515.

Hayes, M. J., and Decker, W. L., 1996, Using NOAA AVHRR data to estimate maize production in the United States Corn Belt. *International Journal of Remote Sensing*, **17**, 3189–3200.

Hodges, T., Botner, D., Sakamoto, C., and Hang, J., 1987, Using the Ceres-maize model to estimate production for the US Cornbelt. *Agriculture and Forest Meteorology*, **40**, 293–303.

IBGE, 1999, Anuário Estatístico do Brasil. *Fundação Instituto Brasiliro de Geografia e Estatística*, Rio de Janeiro, Brasil. http://www.sidra.ibge.gov.br/bda/agric.

Kidwell, K. B., 1996, *NOAA Polar Orbiter Data Users Guide* (Washington, DC, USA: Satellite Data Service Division, NESDIS/NOAA).

Kogan, F. N., 1990, Remote sensing of weather impacts on vegetation in non-homogenous areas. *International Journal of Remote Sensing*, **11**, 1405–1419.

Kogan, F. N., 1995, Droughts of the late 1980s in the United States as derived from NOAA polar orbiting satellite data. *Bulletin of the American Meteorological Society*, **76**, 655–668.

Kogan, F. N., 1997, Global drought watch from space. *Bulletin of the American Meteorological Society*, **78**, 621–636.

Liu, W. T., and Kogan, F. N., 1996, Monitoring drought using vegetation condition index. *International Journal of Remote Sensing*, **17**, 2761–2782.

McQuigg, J. D., 1975, *Special Technical Report: Economic impacts for weather variability* (Columbia, MO, USA: Department of Atmospheric Science, University of Missouri).

Quarmby, N. A., Milnes, M., Hindle, T. L., and Silleos, N., 1993, The use of multi-temporal NDVI measurements from AVHRR data for crop yield estimation and prediction. *International Journal of Remote Sensing*, **14**, 199–210.

Rao, C. R. N., and Chen, J., 1995, Inter-satellite calibration linkages for the visible and near-infrared channels of the Advanced Very High Resolution Radiometer on the NOAA-7, -9, and -11 spacecrafts. *International Journal of Remote Sensing*, **16**, 1931–1942.

Rao, C. R. N., and Chen, J., 1996, Post-launch calibration of the visible and near-IR channels of the AVHRR on the NOAA-14 spacecraft. *International Journal of Remote Sensing*, **17**, 2743–2747.

Rasmussen, M. S., 1997, Operational yield forecasting using AVHRR NDVI data: prediction of environmental and inter-annual variability. *International Journal of Remote Sensing*, **18**, 1059–1077.

Tucker, C. J., and Sellers, P. J., 1986, Satellite remote sensing of primary production. *International Journal of Remote Sensing*, **7**, 1395–1416.

Unganai, L. S., and Kogan, F. N., 1998, Drought monitoring and corn yield estimation in Southern Africa from AVHRR data, *Remote Sensing of Environment*, **63**, 219–232.

USDA (1994), Major world crop areas and climatic profiles. USDA. *Agricultural Handbook*, No. 664: 157–170.

Van Dijk, A., Callis, S. L., and Sakamoto, C. M., 1987, Smoothing vegetation index profiles: an alternative method for reducing radiometric disturbance in NAA/AVHRR dada. *Journal of Photogrammetric Engineering & Remote Sensing*, **63**, 1059–1067.

Weinreb, M. P., Hamilton, G., and Brown, S., 1990, Nonlinearity correction in calibration of the Advanced Very High Resolution Radiometer infrared channels. *Journal of Geophysical Research*, **95**, 7381–7388.